الجانب اللغوي للمعجم الحاسوبي للغة العربية

د/ محمود إسماعيل صالح
أستاذ اللسانيات التطبيقية ، ج م س

١- مصادر المادة اللغوية للمعجم:

المقدمة:

هناك أساليب معروفة لتحديد مداخل أي معجم لغوي عام ، هي:

أ- نظام التقليبات الذي اتبعه الخليل بن أحمد في معجم العين ، حيث ينظر في الجذور واحتمالات تكوين الألفاظ منها ليترك المهمل منها ويؤخذ بالباقي.

ب- نظام الاشتقاقات من الجذور وفق قواعد الصرف المعروفة ، من ثم حصر ألفاظ اللغة بناء على ذلك .

ت- الاعتماد على المعاجم المتاحة للغة ، قديمها وحديثها ، واستخلاص جميع الألفاظ الواردة فيها .

ث- استخلاص الألفاظ مباشرة من المدونات اللغوية الشاملة التي تمثل اللغة العربية في عصورها المختلفة ، أو في عصر معين كالعصر الحديث مثلا.

في الحالتين الأولى والثانية لابد لنا من التأكد من وجود الكلمة واستعمالها فعليا في اللغة ، وهذا ما يتيحه لنا الأسلوب الثالث ، لأنه مبني على المعاجم أصلا.

ولكن أيا من هذه الأساليب لن يكون شاملا ولن يبين لنا درجات شيوع الألفاظ الحقيقي (وليس الافتراضي) ، ولا مجالات استخدامها الفعلي ، وهو ما ينتج من الأسلوب الرابع ، إن أحسنا اختيار مصادر المدونة وراعينا فيها الشمول من حيث أنواع النصوص ومصادرها و الشمول التاريخي والجغرافي ...إلى غير ذلك .

كذلك يمكننا اعتبار المعاجم العربية المعتمدة مصدرا من مصادر المدونة، فنضمن بذلك اشتمال المدونة على ما ورد في هذه المعاجم .

وأرى أن الأسلوب الأخير هو أفضل هذه الأساليب الأربعة ، للأسباب التي نوردها بعد ، ولذلك سنركز الحديث عنه فيما يلي:

أولا: المدونات:

من المعروف في التوجهات الحديثة في العمل المعجمي أن يستند العمل على المدونات اللغوية المحوسبة التي تمثل اللغة المطلوب إعداد المعجم لها تمثيلا حقيقيا ، لا افتراضيا . وتكمن أهمية المدونة في الخصائص التالية:

أ- الواقعية والتمثيل الحقيقي للغة .

ب- الشمول من حيث المصادر والتنوعات والاستعمالات اللغوية والأساليب والأجناس الأدبية والتخصصات العلمية والتقنية ، وذلك بشرط مراعاة ذلك عند إعداد المدونة .

جـ- إمكانية إخضاعها للتحليل الإحصائي من جوانب مختلفة ولأغراض مختلفة ، مثل التعرف على شيوع الكلمات ، ومصاحباتها اللفظية ، وسياقات استعمالها (من خلال الكشافات السياقية concordance) وغير ذلك من أنواع التحليل الصرفي للغات الاشتقاقية كالعربية .

د- التعرف على شيوع الكلمة وشيوع معانيها المختلفة ونسبة شيوع الكلمة مقارنة بمجموع الكلمات في المدونة ، إضافة إلى شيوعها من عدمه في أنواع النصوص المختلفة ، وهو ما يفيد في استخلاص المصطلحات الشائعة في كل تخصص من التخصصات العلمية والتقنية .

هـ- إمكانية التعرف على شيوع الأوزان والصيغ الصرفية المختلفة .

و- إمكانية إجراء أنواع من التحليل النحوي والتركيبي ، مع توافر بعض المتطلبات اللازمة .

ز- إمكانية إجراء التحليل الصوتي (بوصف الحروف تمثيلا للأصوات في العربية بصورة عامة) للوصول إلى معلومات مختلفة عن الأصوات العربية من حيث شيوعها ومواقعها في الألفاظ إلى غير ذلك .

هناك عدد من المدونات التي تم إعدادها ، ويجري تحديث بعضها، للغة العربية ، منها مدونة المعهد العالي السوري ومدونة صخر ومدونة المعجم المدرسي (من مشروعات مدينة الملك عبد العزيز للعلوم والتقنية) ومدونة شركة آي تي آي ATA التي مقرها لندن . وتعتبر مدونة الشركة المذكورة أكثرها شمولا من حيث تنوع المصادر وحجم المادة (التي تزيد عن مليار كلمة ، كما تذكر الشركة) . كما يجدر بالذكر مجموعات المراجع القاموسية والأدبية والدينية وغيرها مما هو متاح في صورة رقمية ، سواء على شكل أقراص مضغوطة أو على الشبكة العنكبوتية .

هذا ويجب هنا التنويه إلى ملاحظة عدم تكرار المصدر الواحد في المدونة التي تخضع للتحليل ، وهو احتمال يرد عند دمج أكثر من مدونة واحدة .

كيفية الاستفادة من هذه المدونات:

أما كيفية الاستفادة من هذه المدونات فنلخصه فيما يلي:

١- عمل كشافات أو مسادر سياقية concordances من هذه المدونات ، حيث ترد كل كلمة رئيسة مسبوقة ومتبوعة بعدد من الكلمات (وهو مايسمى فنيا KWIC أي الكلمة الأساس في سياق) ، مما يوفر سياقا أدنى لكل كلمة ، مع بيان مصدر الكلمة (أي إحالة إلى النص الكامل الذي وردت فيه) .

٢- الرجوع إلى كامل النص أوالنصوص التي وردت فيها الكلمة المذكورة للتعرف على السياق الأكبر ، مما يعين في تحديد المعنى المقصود من الكلمة.

٣- استخراج الكلمات المصاحبة للكلمات موضع البحث.

٤- استخراج الأمثلة والاستشهادات الخاصة بالألفاظ .

٥- يراعى مفهوم "الكلمة" لتمثل الاسم أو الفعل أو الحرف في صورته المجردة من السوابق واللواحق (وهو مايسمى أحيانا بالجذع – مقابل الجذر) ، سواء أكانت الكلمة جامدة أم مشتقة ، مثل ولد ، والد، مولود ...

٦- القيام بدراسة إحصائية لشيوع الكلمات بالمفهوم المذكور أعلاه ، ثم شيوع معانيها (وفقا للسياقات التي ترد فيها).

ثانيا: المعاجم العربية:

أ- العامة:

حيث إن اللغة العربية تتميز بثباتها النسبي عبر التاريخ ، فإننا نستطيع أن نستفيد من أمهات المعاجم المعروفة كلسان العرب والقاموس المحيط وصحاح الجوهري ، من القديم ، ومعجم ألفاظ القرآن الكريم المعجم الوسيط (من أعمال مجمع القاهرة) والمعجم الأساسي (للألكسو) والمعجم الجديد (التونسي) ، في العصر الحديث .

ولكن لابد لنا من التنبيه إلى أن هذه المعاجم مهما حجمها والجهد المبذول في إعدادها تقصر عن حصر ألفاظ اللغة العربية في العصور المختلفة وما طرأ عليها من تغيرات . ويكفينا أن ننظر إلى ما قام به المستشرق رينهارت دوزي في عمله "تكملة المعاجم العربية" ، وكذلك مئات المعاجم المتخصصة التي صدرت وتصدر في العالم العربي في الآونة الأخيرة.

كيفية الاستفادة من هذه المعاجم:

أما كيفية الاستفادة من هذه المعاجم فنلخصه فيما يلي:

١- التعرف على شيوع الكلمة واشتراك المعاجم المختلفة في إيرادها.
٢- التعرف على المعاني المختلفة للألفاظ .
٣- الاستفادة منها في تعريف الألفاظ والشواهد الواردة لاستعمالاتها ، مع مراعاة العصر الذي ألف فيه كل معجم ، وتلافي العيوب التي تعاني منها بعض هذه المعجمات من الناحية المعجمية الفنية (كالتعريفات غير الدقيقة أو الناقصة علميا أو استخدام ألفاظ غير مألوفة ...)
٤- التعرف على المعاني والاستعمالات المستحدثة ، عن طريق مقارنة المعاجم القديمة بالحديثة.
٥- المعرفة ولو بصورة جزئية بالكلمات المولدة والدخيلة ومصادرها .

ب- الخاصة:

تزخر العربية بعدد كبير من المعاجم المتخصصة ، من أهمها مايلي:

أ- في مجال المعاني :

فقه اللغة وسر العربية للثعالبي ، وجواهر الألفاظ لابن قدامة ، والألفاظ الكتابية للهمداني ، وأدب الكاتب لابن قتيبة ، والمفصل لابن سيده (وصورته المعدلة والمختصرة له بعنوان الإفصاح في فقه اللغة لحسن موسى وعبدالفتاح الصعيدي) ، والفروق في اللغة لأبي هلال العسكري ، والرافد لأمين آل ناصر الدين ، و نجعة الوارد في المترادف والمتوارد لليازجي ، وكنز اللغة العربية (موسوعة في المترادفات والأضداد والتعابير). لحنا غالب ، والمكنز الكبير لأحمد مختار عمر ، والمكنز العربي المعاصر لمحمود صيني وآخرين .

جدير بالذكر أن المعاجم المذكورة تتدرج تاريخيا من القديم إلى الحديث ثم المعاصر. وهذا ويوصى بالبحث عن طبعتي فقه اللغة والإفصاح اللتين بهما كشافات للألفاظ ، للاستفادة منها في بناء معجم المترادفات والمتجانسات.

كيقية الاستفادة منها:

وتتم الاستفادة من هذه المعاجم في مجالات الترادف والأضداد والموضوعات ، مثل أسماء الظواهر الطبيعية أو أسماء الانسان أو الحيوان في أعمار مختلفة أو أسماء الأصوات باختلاف مصادرها أو الإبل أو الخيل بألوانه المختلفة ...إلي غير ذلك .

ب- مجال التعبيرات الاصطلاحية والمصاحبات اللفظية:

المعجم السياقي للتعبيرات الاصطلاحية لمحمود صيني وآخرين ، ومعجم التراكيب والعبارات الاصطلاحية القديم منها والمولد والمؤلف لأحمد أبو سعد ، ومعجم الأفعال المتعدية بحرف لموسى آل حمدي ، معجم الأمثال للميداني ، معجم الأمثال العربية لمحمود صيني وآخرين .

يجدر بنا أن نشير إلى (أ) أن المعاجم العامة ترد بها أحيانا التعبيرات الاصطلاحية ، خاصة في سياق الاستشهاد و (ب) أن عددا لا بأس به من التعبيرات الاصطلاحية مصدرها الأمثال .

الاستفادة من هذه المعاجم:

تستقى من هذه المعاجم التعبيرات الاصطلاحية (أي العبارات التي لا تفهم بمجرد فهم معاني الكلمات التي تتكون منها) وكذلك بعض المعلومات المتعلقة بالمصاحبات اللفظية .

جـ- معاجم المولد والدخيل:

هناك عدد من المعاجم القديمة المشهورة للألفاظ الدخيلة والمعربة ، من أشهرها: المعرَّب من الكلام الأعجمي لأبي منصور الجواليقي ومنارالسبيل فيما في اللغة العربية من الدخيل لمحمد الأمين فضل الله المحبي . ويبدو أن المكتبة العربية الحديثة تفتقر إلى مثل هذه المراجع المهمة ، باستثناء الجهود التي قام بها الدكتور ف عبد الرحيم من الجامعة الإسلامية بالمدينة المنورة . ومن هذه الجهود: القول الأصيل فيما في العربية من الدخيل الذي يشتمل على حوالي خمسمائة كلمة (من منشورات مكتبة لينة بدمنهور ، مصر ، في ١٤١١هـ) وكتابه : سواء السبيل فيما في العربية من الدخيل الذي صدر من مكتبة دار المآثر بالمدينة المنورة في ١٤١٩هـ) ، ويشتمل على ٤٠٠ كلمة إضافية مما في العربية الحديثة من الدخيل فيه بيان مصادرها الأجنبية ومعانيها الأصلية . كما للمؤلف المذكور كتابا ثالثا صدر بعنوان : الدخيل في اللغة العربية الحديثة ولهجاتها ، صدر في حلب في ١٣٩٣هـ .

الاستفادة من هذه المعاجم:

لاشك أن مثل هذه المعاجم ستفيد في التعريف بتاريخ بعض الألفاظ الدخيلة في اللغة العربية في العصور السابقة والحديثة ومصادرها المختلفة.

د- معاجم المصطلحات :

تزخر المكتبة العربية بمعاجم المصطلحات القديمة والحديثة ، ومن أشهرها: مفاتيح العلوم للخوارزمي والتعريفات للجرجاني وكشاف اصطلاحات الفنون للتهانوي ، في القديم . أما في العصر الحديث فهناك آلاف من هذه المعاجم المتخصصة في كل فن وعلم . ولمكتب تنسيق التعريب في الوطن العربي ولمجامع اللغة العربية دور بارز في إعداد عدد كبير من مسارد المصطلحات العلمية والفنية . ومن أشهر المعاجم العلمية والفنية الشمولية الصبغة معجم أحمد الخطيب للمصطلحات العلمية والفنية ومعجم معهد الإنماء العربي الذي كان ترجمة لمعجم McGraw-Hill Dictionary of Science and

Technology. ويجدر بنا هنا التنويه إلى ضرورة حصر المستجدات من المصطلحات ، خاصة في مجال المعلوماتية وعلم الجينات وماشابهها مما قد يكون غائبا في المعجمين المذكورين .

الاستفادة من هذه المعاجم:

يستفاد من هذه المعاجم في التعرف على أهم المصطلحات الفنية في حقول المعرفة المختلفة وتعريفاتها، لإدراج المهم والشائع منها في المعجم الحاسوبي .

هـ- معاجم الأخطاء اللغوية:

عالج المتقدمون قضايا الأخطاء اللغوية ، خاصة المعجمية منها ، في كتب مايعرف ب"لحن العامة" . وفي العصر الحديث نجد عددا من المراجع الخاصة بهذا الموضوع ، منها:
معجم الأخطاء الشائعة ، ومعجم الأغلاط اللغوية المعاصرة لمحمد العدناني ، ومعجم الأخطاء الشائعة : قل ولاتقل لكوكب دياب ، ومعجم الأخطاء الشائعة : تصويبا وشرحا وترجمة لخضر موسى محمد محمود ، ومعجم تصحيح لغة الإعلام لعبدالهادي بو طالب ، وكتاب : ألف خطأ وخطأ لوليد النجار.

الاستفادة من هذه المعاجم:

كما هو واضح من عناوين هذه المراجع ، يمكن الاستفادة منها في التنبيه إلى الأخطاء الشائعة التي يقع فيها المؤلفون والإعلاميون.

ثالثا : كمية المعلومات المستقاة من هذه المصادر:

لاشك أن ذلك يعتمد على حجم المعجم وعدد مداخله والقدر المطلوب إدراجه في المعجم الحاسوبي من عناصر ومواد مختلفة ، فكلما كان المشروع طموحا زادت

جرعة المادة المستقاة من هذه المصادر. كما أنه لابد من مراعاة التناسب بين كم العناصر الأساس في المعجم وبين المعلومات الإضافية أو الألفاظ المتخصصة .

رابعا: أسلوب العمل في حسم أوجه الاختلاف بين مصادر المعلومات وآلية العمل في مراجعة المعلومات:

أرى أنه من الضرورة بمكان تشكيل عدد من اللجان المتخصصة من لغويين وعلماء ومعجميين للاستنارة بآرائهم في حسم أوجه الاختلاف الطارئة .

وتسند كذلك إلى مثل هذه اللجان مهمة مراجعة المعلومات والتثبت من صحتها. كذلك يمكننا الاستفادة من خدمات خبراء متعاونين من شتى البلدان العربية ، يتم التواصل معهم وبينهم بالبريد الإلكتروني .

خامسا : الوسط التي تخزن فيه هذه المعلومات:

من الواضح أن عملا طموحا كالذي بين أيدينا لابد فيه من الاستعانة بالحاسوب في جميع مراحل المشروع ، مثل تخزين المعلومات وتحليلها ومعالجتها بشتى الوسائل ، إضافة إلى توثيق كافة المعلومات توثيقا يمكن الباحث من الرجوع فيه إلى مصدر كل معلومة مهمة في المعجم من خلال قواعد البيانات المختلفة .

سادسا: المقدمة والملاحق:

من المعروف أن المقدمة تكتب بعد إنجاز العمل ليشمل تعريفا به وتوثيقا لما تم في سبيل إعداده وإيضاح أساليب الاستفادة ، إضافة إلى التنويه بالمشاركين في إنجازه إلى غير ذلك من معلومات لا يمكن معرفتها قبل إتمام العمل.

أما الملاحق التي يحال الباحث إليها فأقترح أن تشمل ما يلي:

- ملحقا للمرادفات والمجانسات والأضداد والمقابلات .
- ملحقا للحقول الدلالية أو الموضوعات ومفرداتها (ما يسمى بمعجم المعاني).
- ملحقا للعبارات الاصطلاحية (المسكوكة كما يسميها البعض) .
- ملحقا لتصريفات الأفعال بحسب بابها (مثلا: فتح -يفتَح أو نصر-ينصُر) ونوعها (مثلا: مثال ، أجوف ، معتل ، مضعف ...) .
- ملحقا لجداول الاشتقاق حسب فئة كل مصدر.
- ملحقا بالأخطاء الشائعة في استخدام الألفاظ.
- ملحقا بمصادر الكلمات الدخيل ومعانيها الأصلية.
- ملحقا للبيانات الإحصائية الخاصة بكل جذر وكلمة ترد في مداخل المعجم .
- ملحقا بمصادر النصوص لكل كلمة ، في حال وجود مدونات تشتمل عليها ، مع بيان مواقع الكلمة في النصوص (مثلا : الكتاب الفلاني للمؤلف الفلاني ، صفحة كذا).

أما كيفية الإحالة إلى الملاحق فيكون عن طريق الرموز الحرفية أو الرقمية أو باستخدام نظام hypertext أو بهما معا، حيث يضغط الباحث على الكلمة المعنية أو الرمز المعين فيحيله الحاسوب إلى المعلومات المطلوبة . كما يمكن أن يتم العرض وفق قائمة خيارات menu أو حسب صيغ عرض display formats مختلفة ، بحسب رغبة الباحث عن المعلومة .

ملاحظات عامة على المشروع:

أولا: لا أظن أن بالإمكان حصر جميع الألفاظ الممكن اشتقاقها من الجذور المعروفة وبيان معانيها واستعمالاتها ،في المرحلة الأولى ، حيث إن ذلك يتطلب جهدا وأموالا طائلة . لذلك ينصح بأن يتكون المعجم في مرحلته الأولى من حوالي مائة ألف كلمة (جامدة ومشتقة) تكون مداخل له . ويمكن تنميته بحسب الإمكانات المادية والبشرية المتاحة.

ثانيا: أرى من الضروة بمكان تصنيف الألفاظ إلى أقسام كلام أكثر من الثلاثة المعروفة (اسم ، فعل ، حرف) . فهناك حاجة لبيان أكثر تفصيلا للأسماء (مثل الضمائر المختلفة وأسماء الإشارة والظروف وما يقابل الصفات في اللغات الأخرى) إضافة إلى بيان المبني والمعرب وصيغ الجمع لكل مفرد وصيغة الإفراد لأسماء الأجناس وبيان الجنس النحوي ، وللأفعال (من حيث لزومها وتعديها لمفعول أو مفعولين ...إلى غير ذلك) ، مما هو مذكور في المشروع.

ثالثا : بخصوص نطق الألفاظ ، لا أرى ضرورة للجانب التطريزي أو التنغيمي intonation . بل أرى عدم إمكانية ذلك لأن التنغيم ظاهرة تتعلق بالجمل وليس بالكلمات ، كما أنه يخضع لغرض المتحدث من إخبار أو استفهام أو تعجب ...إلى غير ذلك من أغراض.
لكن يمكننا توفير نطق الكلمات والتعبيرات الاصطلاحية من حيث صوامتها وحركاتها ونبرها (مع تسكين أواخرها) ، وذلك من خلال التسجيل الصوتي المضغوط compressed أو الكلام المؤلف أو المركب اصطناعيا synthetic speech ، على أن يراعى اعتبار النطق القرآني معيارا لصحة مخارج الحروف وصفاتها ، وذلك لتجنب مشكلة الاختلافات اللهجية في نطق الكلمات .

رابعا : فيما يخص المصطلحات ، لابد من التنويه إلى التخصص عند ذكر المعنى ، حيث إن الكلمة الواحدة قد تتعدد معانيها وفقا لتخصصات مختلفة.

المعجم الحاسوبي للغة العربية

مقترح بالجانب الحاسوبي للمعجم

د/ إبراهيم الخراشي
مدينة الملك عبد العزيز للعلوم والتقنية ، الرياض

تقدم هذه الوثيقة إجابات أوليه وموجزة على النقاط الخاصة بالجانب الحاسوبي للمعجم ويمكن الاستفاضة فيها بعد التعرف التفصيلي على اللجوانب اللغوية للمعجم

أولاً: لغة البرمجة

يقترح استخدام لغة برمجة تناقلية portable وتعمل على شريحة واسعة من الأجهزة الحاسوبية بما فيها الأجهزة الكفية والجوالات ومختلف النظم التشغيلية. وباستعراض سريع للغات البرمجة يمكن الأخذ بأحد اللغتين التاليتين:

- لغة جافا
- لغة سي ++

ولضمان الانتشار الأوسع والتناقلية للمعجم الحاسوبي عند الانتهاء من تطويره التناقل بين بيئتي التطوير المقترحتين أعلاه مع الحفاظ على أن تكون ترميز البيانات في المعجم واحدة.

ثانياً: الوسط البرمجي لقواعد معطيات المعجم

في بيئة التطوير ورغبة في تسهيل العمل يمكن استخدام أيٍ من أنظمة قواعد البيانات العلائقية Relational Database Systems والتي منها على سبيل المثال ميكروسوفت أكسس MS Accessأو خادم قواعد البيانات من ميكروسوفت MS SQL Server أو ماي أس كيو إل mySQL المجاني

ويفضل عند نشر المعجم للمستفيد النهائي أن تحول بيانات المطلوب نشرها من المعجم إلى ملفات منبسطة flat files وذلك للمزايا التالية:

- تقليل حجم المعجم
- سهولة التثبيت
- زيادة درجة التوافقية مع أنظمة التشغيل المختلفة
- سهولة التحديث للمعجم من قبل المستفيد النهائي

ثالثاً: البرمجيات المساندة لعمل المعجم

يمكن العمل على توفير أو تطوير حزمة من البرمجيات التي تخدم غرض المعجم إما أثناء إعداده أو بعد نشره ومن تلك البرمجيات المساندة ما يلي:

- محلل صرفي
- محلل نحوي
- مشكل آلي
- ...

رابعاً: أسلوب عرض المادة اللغوية

يتيح الحاسوب إمكانات عديدة لعرض المادة اللغوية وتعتمد أساليب العرض المختلفة على نوع المحتوي اللغوي للمعجم بعد وضع التصور النهائي للجانب اللغوي من المعجم. ويمكن سرد العديد من إمكانات العرض المختلفة والتي قد تشمل:

- عرض المادة اللغوية بناءا على الجذر

- عرض المادة اللغوية بمعرفة المصدر /الساق كمدخل
- عرض المادة اللغوية بمعرفة على الكلمة كاملة
- عرض المادة اللغوية حسب سياقها في الذخيرة اللغوية
- عرض المواد اللغوية وربطها بمترادفاتها اللغوية
- عرض المواد اللغوية وربطها بمضاداتها اللغوية
- عرض التعابير المسكوكة
- عرض الإحصاءات المختلفة للغة والمعجم
- العرض حسب التصنيف الشجري للمفردات
- العرض بالدلالة

خامساً: مصادر الوسائط المتعددة

إضافة لمواد الوسائط المتعددة للمتاحة على أقراص مدمجة المتداولة والمتاحة بالمجان أو بمقابل مادي، يفضل الاستعانة بجهة إعلامية لتوفير أو إنشاء مواد تكون أكثر ملائمة لطبيعة المعجم

سادساً: أسلوب عرض الوسائط المتعددة

تعتبر الوسائط المتعددة أسلوب جيد لتقريب مفهوم مداخل المعجم وكما تقدم فإن الحاسوب يتيح إمكانات عديدة لعرض محتوى المعجم بما فيه الوسائط المتعددة المرتبطة بالمداخل وتعتمد أساليب العرض المختلفة على نوع المحتوي اللغوي ويدخل فيه الحس الفني ولهذا فإنه يقترح طلب تقديم مقترح لأساليب العرض من قبل جهة دعائية/إعلامية

سابعاً: خيارات الدخل والخرج

هل لهذا المدخل علاقة بأساليب البحث والعرض ؟ أو أساليب إدخال المادة اللغوية وطرق تحليلها؟

ثامناً: أنواع البحث

مع توفير سبل البحث المختلفة مثل البحث بالجذر أو الساق/المصدرأو الكلمة المجردة أو الكلمة بملحقاتها يمكن توفير سبل بحث مختلفة وتساعد مستخدم المعجم في الاستفادة القصوى من محتوياته ومن الأمثلة على أنواع البحث التي يمكن سردها في هذا السياق مايلي

- الكلمات بمعرفة كلمة وردت في صنف محدد من النصوص (قرأن، شعر، عصر تاريخي محدد)
- البحث بالترادف والتضاد
- البحث بالعلاقات العلوية والدنيا (مثال البحث عن "الحمضيات" ينتج عنه ماله علاقة بهذه المفردة مثل البرتقال والليمون ...)
- الكلمات التي وردت على وزن محدد
- الأسماء المركبة
- أسماء الذوات
- الشيوع للمفردات والجذور والحروف والأوزان واللواصق
- ...

تاسعاً: أنواع الإحصائيات

أنظر ثامناً

عاشراً: أسلوب العمل لتحديث المعجم:

بعد بناء النواة الأولى (النسخة الأولى) للمعجم يقترح وضع أسلوب للتحديث والاقترحات مبني على المشاركة الجماعية من خلال شبكة الإنتنرنت مع مركزية التحكم في قبول ورفض المادة المقترحة

حادي عشر: أسلوب التعامل مع المعجم

يمكن طرح المعجم عبر عدة وسائط منها الأقراص المدمجة والإنترنت كما يعتقد أن ينتج من المشروع العديد من الأدوات الخدمية التي قطعاً ستفيد الكتّاب باللغة العربية والتي يمكن التفاعل معها عبر شبكة الإنترنت حيث يمكن لمستخدم الإنترنت وضع النص المطلوب تحليله صرفياً في مربع نصي على الإنترنت ومن ثم يقوم النظام بتحليل النص وإعادته للمستخدم بعد معالجته ومن تلك الأدوات الخدمية ما يلي:

- التحليل الصرفي
- التشكيل الإملائي
- التصحيح اللغوي
- التصحيح النحوي
- التعرف على أنواع المفردات
- التعرف على المسكوكات
- التعرف على أسماء الذوات
- ...